IN THE CLAIMS

This listing of claims will replace all prior versions, and listings, of claims in the application. An identifier indicating the status of each claim is provided.

Listing of Claims

1. (Currently Amended) An audience state estimation system comprising:

imaging device for imaging an audience and generating a video signal relative to the audience thus imaged;

movement amount detection device for detecting a movement amount of said audience based on said video signal,

wherein the movement amount is selected to estimate an audience state based on a contents provision state,

wherein the movement amount detection device discriminates and extracts a pixel range which is a flesh-color area identifying flesh color from said video signal, divides the extracted flesh-color area into blocks, and calculates a movement vector for each of the divided blocks,

wherein the blocks include a face block representing a face unit of the audience and a hand block representing a hand unit of the audience, and block matching of a current image and a next or previous frame image is performed for each of the blocks,

wherein the movement vector is the movement direction and the movement amount when a result of the block matching indicating images of the blocks are matched,

Frommer Lawrence & Haug LLP
745 Fifth Avenue
New York, NY 10151
212-588-0800
Customer Number 20999

3 of 20

00939180.DOC

wherein each of the divided blocks includes a plurality of pixels, and each of the plurality of pixels identifies flesh color; and

estimation device for estimating an audience state based on a comparison result of said movement amount and a predetermined reference level.

2. (Original) The audience state estimation system according to claim 1, wherein said movement amount detection device determines movement vectors of the imaged audience based on said video signal, and wherein an average movement amount showing an average of magnitudes of the movement vectors is set as the movement amount of said audience.

3. (Canceled)

4. (Original) The audience state estimation system according to claim 1, wherein said movement amount detection device determines movement vectors of the imaged audience based on said video signal and calculates an average movement amount showing an average of magnitudes of the movement vectors, and wherein a time macro movement amount is set as the movement amount of said audience, said time macro movement amount being an average of the average movement amounts in a time direction thereof.

5. (Previously Presented) The audience state estimation system according to claim 1, wherein when said movement amount is larger than the predetermined level, said

estimation device estimates said audience state to be in any one of states of beating time with the

hands and of clapping.

6. (Currently Amended) An audience state estimation system comprising:

imaging device for imaging an audience and generating a video signal relative to

the audience thus imaged;

movement periodicity detection device for detecting movement periodicity of said

audience based on said video signal,

wherein the movement periodicity is selected to estimate an audience state

based on a contents provision state,

wherein the movement periodicity detection device discriminates and

extracts a pixel range which is a flesh-color area identifying flesh color from said video signal,

divides the extracted flesh-color area into blocks, and calculates a movement vector for each of

the divided blocks ,

wherein the blocks include a face block representing a face unit of

the audience and a hand block representing a hand unit of the audience, and block matching of a

current image and a next or previous frame image is performed for each of the blocks,

wherein the movement vector is the movement direction and the

movement amount when a result of the block matching indicating images of the blocks are

matched,

wherein each of the divided blocks includes a plurality of pixels,

and each of the plurality of pixels identifies flesh color; and

estimation device for estimating an audience state based on a comparison result of

the movement periodicity of said audience and a predetermined reference level.

7. (Original) The audience state estimation system according to claim 6, wherein

said movement periodicity detection device determines movement vectors of the imaged

audience based on said video signal, calculates an average movement amount showing an

average of magnitudes of the movement vectors, and detects an autocorrelation maximum

position of the average movement amount, and wherein variance of the autocorrelation maximum

position is set as said movement periodicity.

8. (Original) The audience state estimation system according to claim 7, wherein

the variance is calculated using a signal in a frame range, said frame range being decided on the

basis of the periodicity of said audience state to be estimated.

9. (Previously Presented) The audience state estimation system according to

claim 6, wherein a ratio of low-frequency component in the average movement amount is set as

said movement periodicity.

10. (Original) The audience state estimation system according to claim 9, wherein

a frequency range of the low-frequency component is decided according to the periodicity of the

said average movement amount transformed to a frequency region to be detected.

11. (Previously Presented) The audience state estimation system according to claim 6, wherein said estimation device estimates said audience state to be in a state of beating time with the hands when said movement periodicity is larger than the predetermined level, and estimates said audience state to be in a state of clapping when said movement periodicity is not larger than said predetermined level.

12-28. (Canceled)

29. (Currently Amended) An audience state estimation system comprising:

input device for inputting and generating at least one of video signal obtained by imaging an audience and audio signal obtained according to sound from said audience;

characteristic amount detection device for detecting, based on said video signal, at least one of a movement amount and movement periodicity of said audience, and for detecting, based on said audio signal, a piece of information on at least one of a volume of sound from said audience, periodicity of said sound, and a frequency component of said sound,

wherein the at least one of a movement amount and movement periodicity is selected to estimate an audience state based on a contents provision state,

wherein the characteristic amount detection device discriminates and extracts a pixel range which is a flesh-color area identifying flesh color from said video signal, divides the extracted flesh-color area into blocks, and calculates a movement vector for each of the divided blocks,

wherein the blocks include a face block representing a face unit of the audience and a hand block representing a hand unit of the audience, and block matching of a current image and a next or previous frame image is performed for each of the blocks,

wherein the movement vector is the movement direction and the movement amount when a result of the block matching indicating images of the blocks are matched,

wherein each of the divided blocks includes a plurality of pixels, and each of the plurality of pixels identifies flesh color; and

estimation device for estimating an audience state based on a comparison result of the detected result of said characteristic amount detection device and a predetermined reference level.

30. (Original) The audience state estimation system according to claim 29, wherein said sound from the audience includes voice.

31. (Currently Amended) An audience state estimation method comprising:

imaging an audience and generating a video signal relative to the audience thus imaged;

detecting a movement amount of said audience based on said video signal,

wherein the movement amount is selected to estimate an audience state based on a contents provision state,

discriminating and extracting a pixel range which is a flesh-color area identifying

flesh color from said video signal;

dividing the extracted flesh-color area into blocks;

calculating a movement vector for each of the divided blocks ,

wherein the blocks include a face block representing a face unit of the

audience and a hand block representing a hand unit of the audience, and block matching of a

current image and a next or previous frame image is performed for each of the blocks,

wherein the movement vector is the movement direction and the

movement amount when a result of the block matching indicating images of the blocks are

matched,

wherein each of the divided blocks includes a plurality of pixels, and each

of the plurality of pixels identifies flesh color; and

estimating an audience state based on a comparison result of said movement

amount and a predetermined reference level.

32. (Original) The audience state estimation method according to claim 31,

wherein movement vectors of the imaged audience are determined on the basis of said video

signal, and wherein an average movement amount showing an average of magnitudes of the

movement vectors is set as the movement amount of said audience.

33. (Original) The audience state estimation method according to claim 31,

wherein movement vectors of the imaged audience are determined based on said video signal,

and an average movement amount showing an average of magnitudes of the movement vectors is

calculated, and wherein a time macro movement amount is set as the movement amount of said

audience, said time macro movement amount being an average of the average movement

amounts in the time direction thereof.

34. (Previously Presented) The audience state estimation method according to

claim 31, wherein when said movement amount is larger than the predetermined level, said

audience state is estimated to be in any one of states of beating time with the hands and of

clapping.

35. (Currently Amended) An audience state estimation method comprising:

imaging an audience and generating a video signal relative to the audience thus

imaged;

detecting movement periodicity of said audience based on said video signal,

wherein the movement periodicity is selected to estimate an audience state

based on a contents provision state,

discriminating and extracting a pixel range which is a flesh-color area identifying

flesh color from said video signal;

dividing the extracted flesh-color area into blocks;

calculating a movement vector for each of the divided blocks ,

wherein the blocks include a face block representing a face unit of the audience and a hand block representing a hand unit of the audience, and block matching of a current image and a next or previous frame image is performed for each of the blocks,

wherein the movement vector is the movement direction and the movement amount when a result of the block matching indicating images of the blocks are matched,

wherein each of the divided blocks includes a plurality of pixels, and each of the plurality of pixels identifies flesh color; and

estimating an audience state based on a comparison result of the movement periodicity of said audience and a predetermined reference level.

36. (Original) The audience state estimation method according to claim 35, wherein movement vectors of the imaged audience are determined on the basis of said video signal, an average movement amount showing an average of magnitudes of the movement vectors is calculated, and an autocorrelation maximum position of the average movement amount is detected, and wherein variance of the autocorrelation maximum position is set as the movement periodicity.

37. (Previously Presented) The audience state estimation method according to claim 35, wherein a ratio of low-frequency component in the average movement amount is set as said movement periodicity.

38. (Previously Presented) The audience state estimation method according to claim 35, wherein when said movement periodicity is larger than the predetermined level, said audience state is estimated to be in a state of beating time with the hands, and when said movement periodicity is not larger than said predetermined level, said audience state is estimated to be in a state of clapping.

39-54. (Canceled)

55. (Currently Amended) An audience state estimation method comprising:

generating any one of a video signal obtained by imaging an audience and an audio signal according to sound from said audience;

detecting, based on said video signal, at least one of a movement amount and movement periodicity of said audience,

wherein the at least one of a movement amount and movement periodicity is selected to estimate an audience state based on a contents provision state,

discriminating and extracting a pixel range which is a flesh-color area identifying flesh color from said video signal;

dividing the extracted flesh-color area into blocks;

calculating a movement vector for each of the divided blocks ,

wherein the blocks include a face block representing a face unit of the audience and a hand block representing a hand unit of the audience, and block matching of a current image and a next or previous frame image is performed for each of the blocks,

wherein the movement vector is the movement direction and the

movement amount when a result of the block matching indicating images of the blocks are

matched,

wherein each of the divided blocks includes a plurality of pixels, and each

of the plurality of pixels identifies flesh color;

detecting, based on said audio signal, a piece of information on at least one of a

volume of sound from said audience, periodicity of said sound, and a frequency component of

said sound; and

estimating an audience state based on a comparison result of said detected result

and a predetermined reference level.

56. (Original)   The audience state estimation method according to claim 55,

wherein said sound from the audience includes voice.

57. (Currently Amended)   A non-transitory computer-readable medium storing an

audience state estimation program, executed by a computer- processor, for estimating an

audience state by processing information, said program comprising:

a step of performing any one of detection, based on said video signal obtained by

imaging the audience, for at least one of a movement amount and movement periodicity of said

audience, and detection, based on said audio signal according to sound from said audience, for a

piece of information on at least one of a volume of sound from said audience, periodicity of said

sound, and a frequency component of said sound,

wherein the at least one of a movement amount and movement periodicity

is selected to estimate an audience state based on a contents provision state,

wherein the step of performing detection discriminates and extracts a pixel

range which is a flesh-color area identifying flesh color from said video signal, divides the

extracted flesh-color area into blocks, and calculates a movement vector for each of the divided

blocks ,

wherein the blocks include a face block representing a face unit of

the audience and a hand block representing a hand unit of the audience, and block matching of a

current image and a next or previous frame image is performed for each of the blocks,

wherein the movement vector is the movement direction and the

movement amount when a result of the block matching indicating images of the blocks are

matched,

wherein each of the divided blocks includes a plurality of pixels,

and each of the plurality of pixels identifies flesh color; and

a step of estimating the audience state based on a comparison result of said

detected result and a predetermined reference level.

58. (Previously Presented) The program according to claim 57, wherein said

sound from the audience includes voice.

THE REMAINDER OF THIS PAGE INTENTIONALLY LEFT BLANK